15.6 A 7mW-to-183mW Dynamic Quality-Scalable H.264 Video Encoder Chip

Hsiu-Cheng Chang[1], Jia-Wei Chen[1], Ching-Lung Su[2], Yao-Chang Yang[1], Yao Li[1], Chun-Hao Chang[1], Ze-Min Chen[2], Wei-Sen Yang[2], Chien-Chang Lin[1], Ching-Wen Chen[1], Jinn-Shan Wang[1], Jiun-In Guo[1]

[1]National Chung-Cheng University, Chia-Yi, Taiwan
[2]National Yun-Lin University of Science Technology, Yun-Lin, Taiwan

Improving the hardware efficiency of video coding LSI like MPEG-4/H.264 is a recent design trend in implementing multimedia systems aimed at high-throughput design for high definition (HD) video [1] and low-power design for portable video [2]. However, it is difficult to trade-off power consumption and video quality like programmable processors do for power-adaptive applications. Multimedia processors, though providing quality scalability, still have performance limitations up to D1 video encoding [3]. To achieve high throughput rates, low-power consumption, and flexible quality scalability, a dynamic quality-scalable H.264 video encoder is presented that supports both versatile video resolutions from CIF to HD720 and versatile video qualities on the same video (e.g. D1 video) when operating at different clock rates. Compared to state-of-the-art designs [1], the 470kgates/13.3KB SRAM H.264 video encoder achieves a 49% reduction in gate-count and 61% reduction in internal memory. Moreover, it provides scalable video encoding of 7mW-to-25mW for CIF, 27mW-to-163mW for D1, and 122mW-to-183mW for HD720 with different quality modes.

Figure 15.6.1 shows the block diagram of the design with four MacroBlock (MB) pipelined stages. To reduce hardware complexity, the major functional modules are optimized, including Motion Estimation (ME), Intra-Coding, In-Loop Filter, and Entropy Coding through low complexity algorithms. To support H.264 video encoding with scalable quality, configurations in both the ME and Intra-Coding algorithms are provided through the System Controller. To improve video quality, a pipelined MB-level BU-based rate control scheme is used. To reduce internal memory size, a Predictive Data Store Buffer (PDSB) [4] controller is adopted to efficiently access the intermediate data for video encoding through AHB-based SDR memory.

Figure 15.6.2 shows the scalable algorithm and architecture for ME, including Integer ME (IME) and Fractional ME (FME). The IME algorithm consists of two stages. The IME stage 1 performs block-matching operations on the variable-rate (determined by DSR) 2-D down-sampled search points in the search window to select variable numbers (determined by CN) of good candidates for next stage operations. The IME stage 2 performs local variable-range (determined by LFSR) full-search block-matching operations on the search points around the selected candidates to determine the best motion vectors for all the 41 modes in IME. By setting the variables DSR, CN, and LFSR, the design provides four quality-scalable modes (i.e. QS0, QS1, QS2, and QS3) in doing IME operations. To decide the best mode for encoding the MB, the FME operations are applied to the 41 modes of IME motion vectors separated in two clusters, i.e. Cluster 1 (16×16, 16×8, 8×16, and 8×8) and Cluster 2 (8×4, 4×8, and 4×4). The scalable qualities are provided by selecting different numbers of clusters for doing FME. In the FME algorithm, a Block Size Trend Prediction scheme is adopted to skip the unnecessary FME operations on the IME modes of Cluster 2 if the IME cost in 16×16 mode is lower than that in 8×8 mode. Compared to JM93 [5], the ME design reduces 53%, 73%, 81%, and 87% of the complexity at a cost of 0.13dB, 0.16dB, 0.18dB, and 0.23dB of PSNR loss when operating at QS0, QS1, QS2, and QS3, respectively.

Figure 15.6.3 shows the scalable intra-coding algorithm and architecture. By exploiting spatial correlation of texture between current and neighboring blocks, *CC-SA* searches fewer prediction modes to reduce complexity by 45%. In addition, *PCC-SA* exploits the statistics of intra-coding modes to only search high probability modes for reducing complexity up to 57%. Moreover, *QMB-SA* is proposed to reduce 75% of the complexity in intra-coding on chrominance pixels by observing that human eyes are less sensitive to errors of chrominance pixels than luminance ones. Exploiting the proposed *CC-SA*, *PCC-SA*, and *QMB-SA* to perform different numbers of intra-coding modes, quality scalability of QS0/QS1, QS2, and QS3 in intra-coding are provided with less than 0.2dB of PSNR drop to obtain 1.4 and 1.74 times throughput improvement from QS0/QS1 to QS2 and QS0/QS1 to QS3, respectively. In addition, an *ETBAC* technique is proposed to reduce 30% of the complexity without PSNR loss by early stopping of the intra-prediction when the accumulated costs of I4MB and I16MB are both higher than the inter-prediction cost. The complexity reduction in ME and intra-coding leads to a reduction in gate-count of the proposed design.

Figure 15.6.4 shows the performance of the proposed design with four quality-scalable configurations, including the required memory bandwidth for different video formats. It can respectively encode D1 and HD720 videos at clock rates of 30/40/60/96MHz and 72/108MHz, respectively, at different quality levels with less than 0.6dB of PSNR loss on average. The quality-scalability allows the design to adjust its encoding quality by trading off different amounts of power consumption. It achieves different video recording time, with finite battery charge, when used in power-adaptive coding applications like portable digital video recorders. In addition, a pipelined MB-level BU-based rate control scheme is used to increase video quality by good bit allocations in constant bit-rate video encoding, as shown in Fig. 15.6.5. The BU-based rate control in JM93 requires strong data dependency in encoding each MB, which causes the needed real bit sizes and Mean Absolute Difference (MAD) values to be unavailable when generating the quantization parameters in the MB-pipelined H.264 encoders. To solve this problem, two precise prediction models on rate control are used to achieve a smoother coded bit rate and a fewer number of skipped frames for better video quality than the frame-based rate control scheme [5].

Figure 15.6.6 summarizes the chip implementation. The core size is 4.3×4.3mm^2 and includes 470kgates and 13.3KB of internal memory. The chip micrograph is shown in Fig. 15.6.7. The power consumption is 7mW-to-25mW, 27mW-to-163mW, and 122mW-to-183mW for real-time encoding CIF, D1, and HD720 video with different quality levels, respectively. Figure 15.6.6 also compares the design with existing H.264 video encoders [1,6]. Compared to the state-of-the-art [1] targeted at HD720 video, the design achieves a 49% reduction in gate-count and 61% reduction in internal memory. When operating at 10MHz for CIF video encoding, its power consumption is only 7mW, which is comparable to the 5mW reported in the state-of-the-art MPEG-4 encoder [2] for CIF video. Moreover, with quality-scalable flexibility, this design can be applied to power-adaptive video coding applications by trading off video quality and power consumption dynamically.

Acknowledgements:
The authors thank the National Science Council, the Ministry of Economic Affairs, and the National SoC Program of Taiwan for funding the research. We also the thank Chip Implementation Center for supporting chip fabrication.

References:
[1] Y. W. Huang, et. al., "A 1.3TOPS H.264/AVC Single-chip Encoder for HDTV Applications," *ISSCC Dig. Tech. Papers*, pp. 128-130, Feb., 2005.
[2] C. P. Lin, et. al., "A 5mW MPEG4 SP Encoder with 2D Bandwidth-Sharing Motion Estimation for Mobile Applications," *ISSCC Dig. Tech. Papers*, pp. 412-413, Feb., 2006.
[3] "H.264 software IP suite for DSP C64xx," from ATEME, http://www.ateme.com.
[4] C. C. Lin, et. al., "A 160kGate 4.5kB SRAM H.264 Video Decoder for HDTV Applications," *ISSCC Dig. Tech. Papers*, pp. 406-407, Feb., 2006.
[5] Joint Video Team (JVT) of ISO/IEC MPEG&ITU-T VCEG, "ISO/IEC 14496-10," 2003.
[6] "A Fully Hardwired H.264 Encoder IP," from MMChips: MM5010, http://www.mmchips.com/ product_mm5010.html.

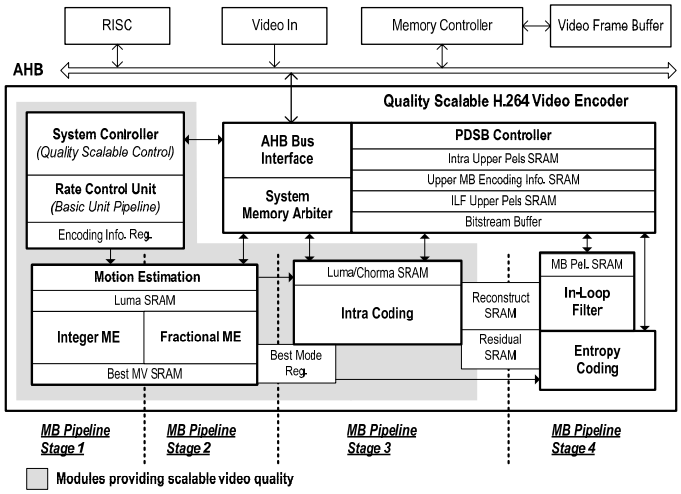


Figure 15.6.1: Block diagram of quality-scalable H.264 video encoder.

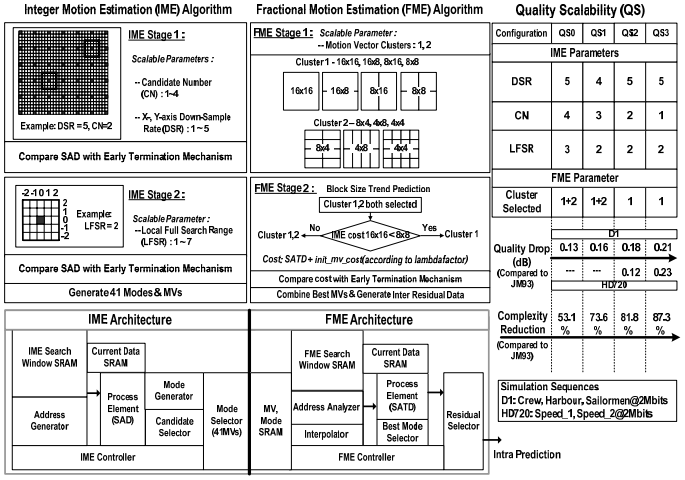


Figure 15.6.2: Scalable motion estimation algorithm and architecture.

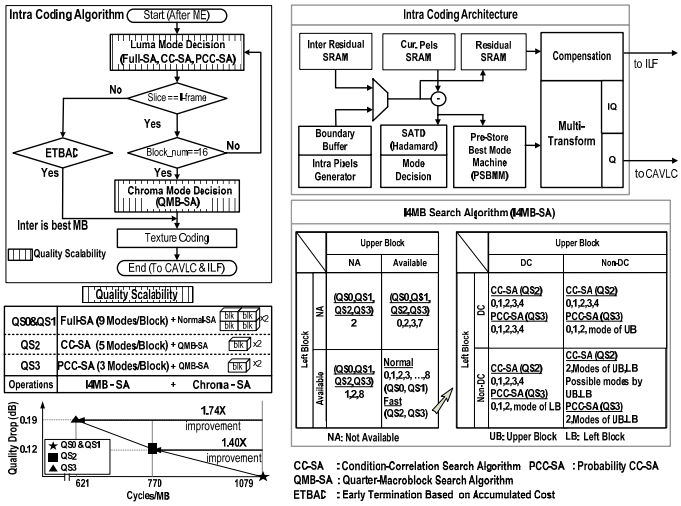


Figure 15.6.3: Scalable intra coding algorithm and architecture.

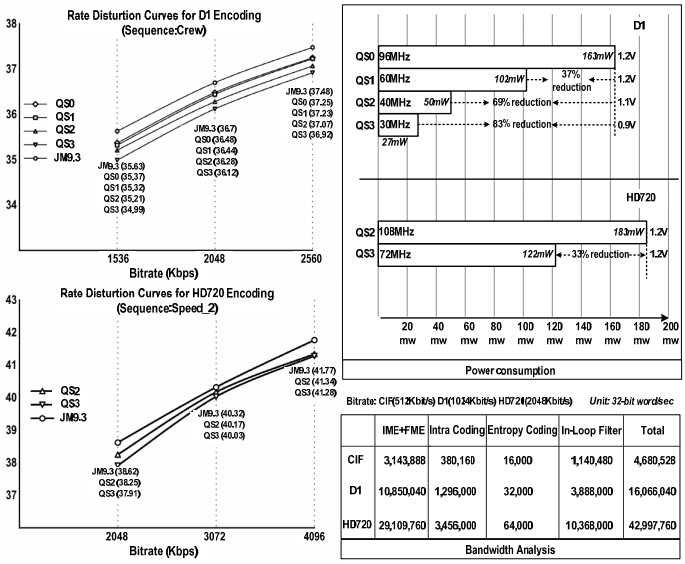


Figure 15.6.4: Quality-scalable modes in the proposed H.264 video encoder.

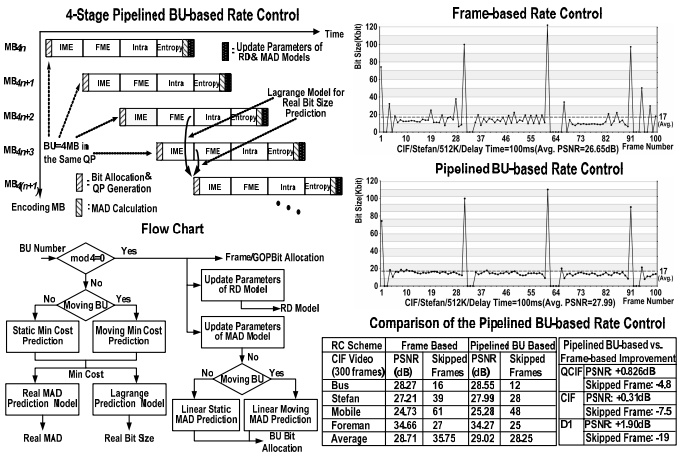


Figure 15.6.5: Pipelined BU-based rate control scheme for improving video encoding quality.

	This Work	ISSCC 2005 [1]	MMChips [6]	ISSCC 2006 [2]
Maximum Resolution	1280 x720 (HD720)	1280 x 720 (HD 720)	720 x 480 (D1)	640 x480 (VGA)
Profile	H.264 Baseline	H.264 Baseline	H.264 Baseline	MPEG -4 Simple Profile
Gate Count	470 K	922 .8K	500 K	201 K
Internal Memory	13 .3KB	34 .72KB	28KB	4.56KB
Technology	130 nm	180 nm	---	180 nm
Operating Frequency (30 fps)	72 /108 MHz for HD720 30 /40 /60 /96 MHz for D1 10 /12 /18 /28 MHz for CIF	108 MHz for HD720 81 MHz for D1	54 MHz for D1	28 .5 MHz for VGA 9 .5MHz for CIF
Power Consumption	183 mW @ 108 MHz /1 .2V 27 mW @ 30MHz /0 .9V 7mW @ 10 MHz /0 .7 V	785 mW @ 108 MHz 581 mW @ 81 MHz	---	18 mW @ 28 .5 MHz /1 .4V 5mW @9 .5MHz /1 .3 V

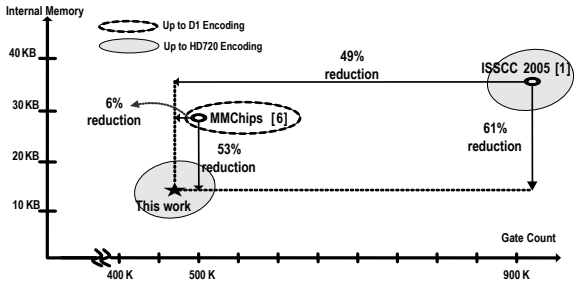
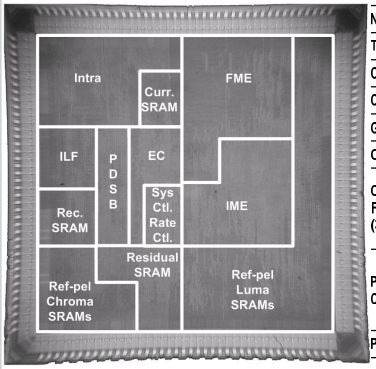


Figure 15.6.6: Comparison with the state-of-the-art H.264 encoders.

Continued on Page 603



Name	Quality Scalable H.264 Video Encoder
Technology	TSMC 0.13μm 1P8M CMOS
Chip Size (Including PAD)	4.9 x 4.9 mm²
Core Size	4.3 x 4.3 mm²
Gate Count	470K
On-Chip Memory	13.3K Bytes (SRAM)
Operating Frequency (30fps)	72/108 MHz for HD720 30/40/60/96 MHz for D1 10/12/18/28 MHz for CIF
Power Consumption	183mW@108MHz/1.2V for HD720 27mW@30MHz/0.9V for D1 7mW@10MHz/0.7V for CIF
Package	160CQFP

Figure 15.6.7: Chip micrograph and specifications.

